2014-12-01

S-CAPE Testing for Higher Proficiency Levels and Other Factors That Influence Placement at Brigham Young University

Elizabeth Robinson
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd
Part of the Spanish and Portuguese Language and Literature Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/4354

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
S-CAPE Testing for Higher Proficiency Levels
and Other Factors That Influence Placement
at Brigham Young University

Elizabeth Robinson

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Arts

Cherice Montgomery, Chair
Troy L. Cox
Nieves Perez Knapp
Gregory Thompson

Department of Spanish and Portuguese
Brigham Young University
December 2014

Copyright © 2014 Elizabeth Robinson
All Rights Reserved
ABSTRACT

S-CAPE Testing for Higher Proficiency Levels and Other Factors That Influence Placement at Brigham Young University

Elizabeth Robinson
Department of Spanish and Portuguese, BYU
Master of Arts

Brigham Young University (BYU) first implemented the Spanish Computer Adaptive Placement Examination (S-CAPE) during the Fall Semester of 1986 and it has been used ever since. The S-CAPE was designed to determine course placement into beginning and intermediate classes for students who have previously studied Spanish. A 10% increase occurred this year (2014) in students who have served missions for The Church of Jesus Christ of Latter-day Saints. Many of these returned missionaries gained language proficiency on their missions, and some go to BYU to begin or continue their studies. Because of the increase in enrollment of students with intermediate and advanced Spanish fluency, the BYU Department of Spanish and Portuguese needed a way to accurately place these students. This study analyzed the S-CAPE to see if it was reliable and capable of placing more advanced students. The S-CAPE was not originally designed to place students above SPAN 206. In addition, other factors that contribute to student placement at BYU are evaluated. Recommendations are made for improving the validity of the S-CAPE, as well as the language skills tested by the S-CAPE. Further recommendations are made to upgrade the process of placing students registering for Spanish at BYU.

Keywords: placement test, test reliability and validation, advanced-level proficiency, Spanish language, returned missionaries
ACKNOWLEDGMENTS

To begin, I would like to thank the members of my thesis committee—Dr. Nieves Knapp, Dr. Gregory Thompson, and especially Dr. Cherice Montgomery—who willingly spent many extra hours helping me organize and edit my thoughts. I would also like to thank Dr. Troy Cox for joining my committee and helping with much of the quantitative analysis of this thesis. A professor who was never an actual member of my committee, but who spent many hours working with me and who was also the author of the original S-CAPE exam was Dr. Jerry Larson. Next, I would like to thank Bob Bockholt and Russell Hansen for administering the S-CAPE and for collecting S-CAPE and survey data. Additionally, I would like to thank others involved in the process: Rick Winterton, Tyler Hickok, Rebecca Brazzale, and also my family—especially my parents, Stephen and Susan Robinson, and my brother, Spencer Robinson.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER 1</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>4</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>5</td>
</tr>
<tr>
<td>Research Questions</td>
<td>5</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td>7</td>
</tr>
<tr>
<td>Review of Relevant Literature</td>
<td>7</td>
</tr>
<tr>
<td>Complexities of Placing Language Students</td>
<td>7</td>
</tr>
<tr>
<td>Placement Tests</td>
<td>9</td>
</tr>
<tr>
<td>S-CAPE</td>
<td>10</td>
</tr>
<tr>
<td>Responses to the S-CAPE</td>
<td>20</td>
</tr>
<tr>
<td>Testing for Intermediate and Advanced Proficiency Students</td>
<td>20</td>
</tr>
<tr>
<td>Research Questions</td>
<td>21</td>
</tr>
<tr>
<td>Chapter Summary</td>
<td>21</td>
</tr>
<tr>
<td>CHAPTER 3</td>
<td>23</td>
</tr>
<tr>
<td>Research Design and Methods</td>
<td>23</td>
</tr>
</tbody>
</table>
### Methodology .......................................................... 23

### Participants ...................................................................... 25

### Data Instruments ................................................................ 26

### Data Analysis ...................................................................... 27

### Chapter Summary .......................................................... 29

### CHAPTER 4 ........................................................................ 30

### S-CAPE Findings ............................................................ 30

#### Reliability of the S-CAPE .................................................. 31

#### The S-CAPE and Placement ............................................... 34

#### Analysis of Perceived S-CAPE Validity ............................... 37

#### Survey of the correct placement of the S-CAPE according to class instructors ...................... 37

#### Survey of the effectiveness of the S-CAPE according to students ................................. 38

#### Factors that Changed the Functionality of the S-CAPE ................................................. 39

#### How students placed themselves ...................................... 40

#### Limitations of This Study .................................................. 42

### Chapter Summary .......................................................... 43

### CHAPTER 5 ........................................................................ 44

### Discussion and Implications ............................................... 44

#### The Reliability of the S-CAPE ............................................. 44

#### Teacher and Student Confidence in the S-CAPE ......................................................... 44

#### How Students Place Themselves ........................................... 46
LIST OF TABLES

Table 1 Reliability and Point Biserial Coefficients for the Five Test Forms .................................................. 12
Table 2 Demographic of Students Participating in the S-CAPE Study .................................................................. 255
Table 3 Demographic of Instructors Participating in the S-CAPE Study ............................................................. 266
Table 4 ANOVA: Single Factor Summary for SPAN 105–321 ..................................................................... 355
Table 5 Current and Recommended Placement Scores for the S-CAPE .............................................................. Error! Bookmark not defined.3
# LIST OF FIGURES

(Some images are presented in color.)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Introduction to the Student for the S-CAPE</td>
<td>14</td>
</tr>
<tr>
<td>2.2</td>
<td>Practice Question at the Beginning of the S-CAPE</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>S-CAPE Vocabulary Sample Question 1</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>S-CAPE Vocabulary Sample Question 2</td>
<td>16</td>
</tr>
<tr>
<td>2.5</td>
<td>S-CAPE Grammar Sample Question 1</td>
<td>16</td>
</tr>
<tr>
<td>2.6</td>
<td>S-CAPE Grammar Sample Question 2</td>
<td>17</td>
</tr>
<tr>
<td>2.7</td>
<td>S-CAPE Reading Comprehension Sample Question 1</td>
<td>17</td>
</tr>
<tr>
<td>2.8</td>
<td>S-CAPE Reading Comprehension Sample Question 2</td>
<td>18</td>
</tr>
<tr>
<td>2.9</td>
<td>Sample S-CAPE Administrator’s Report for an Individual Student</td>
<td>19</td>
</tr>
<tr>
<td>4.1</td>
<td>Distribution of the Mean of S-CAPE Scores Divided by 100-, 200- and 300-level Courses</td>
<td>31</td>
</tr>
<tr>
<td>4.2</td>
<td>Results for Test-Retest Interrater Reliability</td>
<td>32</td>
</tr>
<tr>
<td>4.3</td>
<td>95% Confidence Interval Based on the Mean and Standard Deviation of Students in SPAN 105–321</td>
<td>34</td>
</tr>
<tr>
<td>4.4</td>
<td>How Students Felt About the Effectiveness of the S-CAPE as a Placement Tool</td>
<td>39</td>
</tr>
<tr>
<td>4.5</td>
<td>How Students Chose Their Current Class; Survey Options</td>
<td>40</td>
</tr>
<tr>
<td>4.6</td>
<td>Results of the Other Category</td>
<td>42</td>
</tr>
<tr>
<td>5.1</td>
<td>How Student in SPAN 105–321 Placed Themselves</td>
<td>46</td>
</tr>
<tr>
<td>5.2</td>
<td>Course Information for SPAN 205</td>
<td>48</td>
</tr>
<tr>
<td>5.3</td>
<td>Course Information for SPAN 321</td>
<td>49</td>
</tr>
<tr>
<td>5.4</td>
<td>Example S-CAPE Question that Needs Updating</td>
<td>51</td>
</tr>
<tr>
<td>5.5</td>
<td>Analysis of Learning Outcomes for SPAN 101–206 Using Wordle.net</td>
<td>56</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

When I first entered Brigham Young University (BYU) in Fall Semester of 2003 as an undergraduate, I took the S-CAPE (Spanish Computer Adaptive Placement Exam) to decide on the right Spanish class for me. It placed me into SPAN 202 (a 4th semester intermediate Spanish course at that time). I did register for and take that class. However, I felt that conversationally, I would have benefited by starting with an earlier course. Eleven years later, as a student instructor of the beginning Spanish classes at BYU, I saw that my students lacked confidence about what class they should be taking. These experiences led me to examine the placement test and its effectiveness in placing students in an appropriate level class.

In 1984, BYU began developing the S-CAPE to place students with previous Spanish experience into introductory and intermediate Spanish courses. After it was first implemented as the primary means of placement during Fall Semester of 1986, many other institutions adopted the exam. After successful development in Spanish, it was developed in four additional languages—French, German, Russian, and ESL—and is now administered to hundreds of thousands of students each year. BYU has used this placement test to place students for the last 28 years. However, during that time it has not been updated.

All placement tests depend on the context in which they are administered to be effective. Brown and Hudson (2002) stated, “No test—norm-referenced or criterion-referenced—is effective, reliable, and valid unto itself. Tests are effective, reliable, and valid for particular purposes with specific types of students who have a particular range of abilities” (p. 32). Given the fact that the S-CAPE was originally designed and calibrated for the student body of 1986, an evaluation of its effectiveness was necessary.
The demographic of students enrolling in Spanish courses has changed since the S-CAPE was first implemented. A large number of students who take introductory Spanish classes have had previous experience in the language, whether they learned it in school, as heritage learners (students who learn at home from native speakers of the language, which is different from the dominant language of the society), or on a mission. BYU is unique in that many of its students serve missions for The Church of Jesus Christ of Latter-day Saints. As part of this service, many gain language skills, but because some travel to foreign countries and others develop their language skills while serving target language-speaking populations in different areas of the United States, they return with varying proficiencies. Young men serve for two years. Young women serve for a year and a half. Health or other problems for either group may cause them to end their missions earlier than planned, which has an obvious effect on fluency due to a decreased amount of time in the country or using the target language. Missionaries assigned to Spanish-speaking populations receive language and theological training for five weeks in a Missionary Training Center. If they have previous experience with the language, they may only be there for two weeks, depending on their level of fluency in the language. Since missionaries are occasionally sent to areas where multiple languages are spoken, some must acquire language skills during their volunteer experience while working with a certain linguistic community. In October 2012, the missionary age for young men who serve was lowered from 19 to 18, and the age was lowered for young women from 21 to 19. When the announcement was made, many students deferred their enrollment to serve missions and they are expected to return to BYU over the next year. Since the announcement, there has already been a 10% increase in returned missionaries for all of BYU, meaning that now 56% of students are returned missionaries (Hollingshead, 2014). The BYU Department of Spanish and Portuguese recognizes that this
change will magnify the already complicated issue of student placement, especially for returned missionaries. They also anticipate that fewer students will register for entry level classes due to the fact they may return from missions with language proficiency. Those who return will need to know which language class to take if they are desirous to continue their study of Spanish.

BYU currently has beginning Spanish classes that include SPAN 105 and 106 (also known as University Spanish 1 and 2), intermediate Spanish classes that include SPAN 205 and 206 (also known as University Spanish 3 and 4) and a more advanced class called SPAN 321 (for third-year Spanish students). (Other more advanced Spanish courses focus on writing and literature. There are also two foundational courses entitled SPAN 101 and 102). When the S-CAPE was implemented 28 years ago, most Spanish-speaking missionaries were coming from an in-country immersion experience. Thus, the SPAN 321 recommendation made sense as the recommended class for all returned missionaries. However, due to the increased immigration of Spanish speakers to the United States, more missionaries are being called to stateside, Spanish-speaking missions. These missionaries typically do not get an equivalent amount of exposure to the language as those who serve in foreign countries. As a result, it is assumed that they do not make the same language gains. When they return from their missions, some returned missionaries doubt that they have the proficiency necessary to be successful at the SPAN 321 level. They are left to decide on their own which Spanish class they should take. Some have placed themselves as low as SPAN 101. Therefore, the context has changed, but institutional recommendations have not been updated.

After the S-CAPE was originally validated, Larson (1986), the test’s author, stated that, “This range of difficulty [referring to the test items] easily spanned the ability level of students in the lower division Spanish courses [at BYU]” (p. 5). However, due to the changing demographic
of the returned missionary population, BYU needs a tool that can effectively place students with more advanced levels of proficiency. The S-CAPE was tested to see if it was capable of placing students at this level. Correctly placing is more important than it is often characterized. James and Templeman (1992) declared, “Although placement testing may not be as high-stakes as admissions testing, its effect is significant. For students it can delay or prolong their studies…inappropriately placed students may jeopardize the quality and integrity of the course sequence. Therefore, accurate…placement testing is crucial” (p. 82). Correctly placing students with advanced level proficiency is important to BYU.

**Statement of the Problem**

Students anecdotally report that the S-CAPE gives them very different scores if they take it more than once. Therefore, one of the first questions to answer about the S-CAPE is whether or not it is reliable. This needs to be answered before the S-CAPE can be used to place students with higher proficiency into more advanced classes. There is a connection between reliability and validity. A test can be reliable and not valid, but it cannot be valid without also being reliable. As Hughes (2003) stated, “The smaller the proportion of misplacements [which represents a test’s reliability], the more valid the test” (p. 72). If results do vary, then receiving a score from the S-CAPE may not always mean that students know how to place themselves after taking the exam.

Tests also need to be updated. Brown and Hudson (2002) stated, “Any test…should be constantly monitored to insure that it effectively fits the purposes of the program and the types of students, especially in terms of their range of abilities” (p. 33). This has not been the case with the S-CAPE, especially in terms of cutoff scores. Cutoff scores for placing students into the correct course were established when the test was first developed 28 years ago. Those cutoff
scores were changed based on expert opinion years later when two more beginning Spanish classes were added. However, these changes were never empirically validated. The current cutoff scores need to reflect the current demographic of BYU students.

The next question is how effectively the S-CAPE places students into upper-level Spanish classes. Currently, it is recommended that returned missionaries enroll in Spanish 321, a third-year Spanish grammar, reading, and culture class; however, as noted earlier, not all mission experiences result in similar language gains whereas in 1986 (when the S-CAPE was originally implemented) this general recommendation would have been more accurate.

The last question is how or if students are actually using the S-CAPE to place themselves and what other factors contribute to the test’s performance.

**Purpose of the Study**

Because of the current problems with placement into the appropriate Spanish courses at BYU, this study analyzes the reliability of the S-CAPE, how effective it is at placing students, and how it is actually used for placement in the BYU Spanish and Portuguese Department. Due to the unique situation that BYU is in, as described above, there is a need for a placement test that will assess students’ abilities at higher levels, as well as a need to examine the policies that guide placement.

**Research Questions**

The following questions need to be answered in order to solve the problem of placing returned missionaries, which may also enable the more effective placement of students with Spanish experience:

1. To what extent does the S-CAPE still function as a reliable assessment instrument?
2. How effective is the S-CAPE for placing students into SPAN 206 and 321 at Brigham Young University?

3. What factors affect the functionality of the S-CAPE in making correct placement decisions?
CHAPTER 2

Review of Relevant Literature

Placement of language students is complicated. This review of literature begins by discussing issues that contribute to its complexity. Next it looks at the need for placement tests and discusses the principles of validity and reliability. It then describes the S-CAPE (Spanish Computer Adaptive Placement Exam) and what others reported about this test. Afterwards, it talks about the difficulty of placing students with intermediate and advanced proficiency.

Some of the following studies review placement tests for languages other than Spanish. Additionally, most current research has dealt more with the placement of ESL students than English speakers learning a foreign language. However, all of the studies were reviewed with the premise that the general principles that underlie placement tests still apply.

Complexities of Placing Language Students

One of the reasons why placement exams are so valuable is because they help address some of the complexities that come with placing students. Placement is difficult because students who are not true beginners, meaning that they have had some contact with the foreign language previously, enter any given university with varying proficiencies. There are many factors that affect the proficiency with which a student comes to a university. All of them are related either to time on task (the amount of overall time the student has worked with or been exposed to the language) and either the student’s efforts or natural abilities in the language. For example, previous educational experience can affect students’ proficiency: how early they began learning the language, the pedagogical and linguistic abilities of their instructors, and the students’ motivation. Proficiency can also be affected by how students grew up: as heritage speakers of the language or in areas where they had exposure to the language. S.S. Robinson, the Language
Program Specialist in the Department of Slavic and East European Languages and Cultures at The Ohio State University stated, “Heritage speakers cover the spectrum in language proficiency. Some say they are fluent and they can’t pass a basic Russian 101 test, while others are basically native speakers. The rest form a continuum in between these extremes. This is one reason why language testing is so important: so students can be placed in classes according to their abilities and not just their ‘background’ which is wildly different” (2014, personal communication). Although he was only referring to heritage speakers, the same principle is true for other language learners, but usually on a smaller spectrum.

Many universities have tried to simplify the issue of placement by recommending that a certain amount of time studying the language should correlate to Spanish courses. However, the number of years studied before entering a university does not guarantee a certain level of proficiency. Aleamoni and Spencer (1968) found that the number of years students studied before they entered the University of Illinois did not correlate to their language abilities. The researchers found that there was questionable correlation between one year of high school and one semester of college. Results may have been affected by the university’s use of the Modern Language Association (MLA) test, which is a proficiency test, but Hagiwara (1983) had similar results at the University of Michigan. He found that it was “…impossible to devise a dependable formula that equates the number of years of high school language study with that of college semesters” (p. 25). Hagiwara reiterates the claim that placement cannot be correlated to the number of years students studied the foreign language in high school. Moreover, the curricula at different institutions varies, so one year of high school at Institution A may correspond to one semester at Institution B, but may be equivalent to only one term at Institution C.
Furthermore, Aleamoni and Spencer (1968) also showed that the number of intervening years between high school and college affected test scores, and therefore, affected placement and proficiency. As a result, more students have been placed into introductory courses without being true beginners, which often gives them an advantage in terms of grades and can be demotivating to those studying a language for the first time.

There is a common solution to the complexities of assigning students to a language course, and that is the use of a placement exam. In summary of the issues described above, and in defense of the development of a placement test, Larson (1991) stated, “Recognizing the rather imprecise nature of placements on this basis [referring to the placement of students according to the number of years they have previously studied], the need for a more exact and objective—yet convenient—measure is desired” (p. 277). This idea led to the development of the S-CAPE.

**Placement Tests**

A language placement test is used for students who have had some introduction to a foreign language. In order to know where they fit in at a certain school or university, their proficiency is assessed to assign them in a language course. Hughes (2003) stated, “Placement tests, as their name suggests, are intended to provide information that will help to place students at the stage (or in the part) of the teaching programme most appropriate to their abilities” (p. 16). This positioning based on ability is the purpose of the S-CAPE. A number of other universities have experimented with offering these types of tests. Their effectiveness depends on how well they fit the needs of the particular university’s curriculum and needs. “The placement tests that are most successful are those constructed for particular situations. They depend on the identification of the key features at different levels of teaching in the institution. . . . The work that goes into their construction is rewarded by the saving in time and effort through accurate
placement” (p. 17). In addition to meeting the university’s specific needs, a placement test also needs to be reliable and valid. Reliability means that the test is consistently giving students the same score. Validity means that the test is actually testing what it says it is testing. There are multiple types of validity, and three will be described here: content, criterion, and face. Content validity means that what is actually being tested is connected to what is being measured. For example, if the S-CAPE is to measure vocabulary knowledge, then it needs to have questions that include vocabulary. Criterion validity is the extent to which the criteria for the test were actually met. In terms of placement exams, its predictive validity is part of its criterion validity. Last, face validity is the extent to which students and instructors view the test as an adequate assessment for the testing context they are in. Messick (as cited in Brown and Abeywickrama, 2010) cautioned that “validity is not an all-or-none proposition. . . . If in your language assessment procedures you can make a point of primarily focusing on content and criterion validity, then you are well on your way to making accurate judgments about the competence of the learners with whom you are working” (p. 36). For this purpose, these two types of validity (content and criterion validity) have been included in this study. Face validity has been included to understand the S-CAPE from the students’ and instructors’ perspectives.

The following section presents the S-CAPE and its key features: what type of test it is, its reliability, and how it was originally validated.

**S-CAPE.** As mentioned previously, the S-CAPE is the placement exam used at BYU. Its development began in 1984 and during Fall Semester of 1986 it was used for the first time as BYU’s primary means of placing students into Spanish courses. The S-CAPE is a computer adaptive placement test, which means that it adjusts the questions it selects depending on students’ responses.
Adaptive testing is based on item response theory (IRT), a statistical theory that assumes that there is an underlying latent ability distribution for any skill. The ability level of each individual taking a given test lies somewhere along the latent ability continuum. An adaptive test is designed to locate an examinee’s ability along the continuum, rather than to determine the degree to which an individuals’ ability differs from the abilities of others taking the same measure, as is the case with conventional tests (Olsen, 1982, p.1). This enables the test to be criterion-referenced as opposed to norm-referenced, meaning that students are being compared to a set of criteria instead of to others who are taking the same test. This is advantageous when aligning the scores of the test with varying courses since the skills of students will vary compared to course requirements that will stay the same. The S-CAPE was initially given on a computer on BYU campus, and it was later put online to make it more accessible to students and to facilitate the registration process. The S-CAPE’s other benefits included reducing testing time, decreasing test boredom, providing instant feedback, the ability of the examinee to set the pace of the test, the decrease in administrators required, and the improvement in test security (Larson, 1991). Creating a placement test is so difficult and time consuming that many universities still have not adopted these advances.

To establish the validity and the reliability of the S-CAPE, a large number of test items were created and calibrated, which covered Spanish 101 and 102, (1st year Spanish classes) and 201 and 202 (2nd year Spanish classes) at BYU; the delineations of the courses at BYU at that time. “Since the focus of the S-CAPE was to be on lower-division courses (i.e., Spanish 101, 102, and 201) items were written to reflect the curricula of these courses” (Larson, 1991, p. 279). (The S-CAPE was later adapted by the Spanish Department faculty to place students into SPAN 101, 102, 105, 106, 205, and 206. SPAN 101 and 102 were added as foundational courses for
true beginners.) Items were created to test grammar, reading, and vocabulary at each level. Out of the 1,443 items that were originally written, 1,100 were approved by the members of the Spanish faculty at BYU, and they were then put into five different test forms. The tests were administered to 199 students, after which both conventional statistical analyses (Hoyt and KR-21 reliability coefficients and the point (PBIS) biserial discrimination coefficient) and the Rasch statistical analyses were used. Table 1 contains the results of those statistical analyses.

Table 1

<table>
<thead>
<tr>
<th>Test Form</th>
<th>Part</th>
<th>Hoyt</th>
<th>KR-21</th>
<th>PBIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>.95</td>
<td>.93</td>
<td>.79</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.94</td>
<td>.92</td>
<td>.82</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>.96</td>
<td>.95</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.95</td>
<td>.94</td>
<td>.82</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>.97</td>
<td>.96</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.97</td>
<td>.96</td>
<td>.82</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>.96</td>
<td>.95</td>
<td>.79</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.95</td>
<td>.94</td>
<td>.83</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>.93</td>
<td>.91</td>
<td>.77</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.87</td>
<td>.83</td>
<td>.85</td>
</tr>
</tbody>
</table>

n = 199

Notice the high level of reliability. Lower numbers are in the PBIS category, which represents the correlation between an item being right or wrong and the participant’s overall score. The Rasch analysis distinguished item difficulty indices that were “‘anchored’ together and located along a single difficulty/ability continuum.” Ninety-six levels of difficulty were originally
recognized, but they were reduced to 51 levels, “covering a difficulty range of -2.0 to +5.8 logits” (Larson, 1986, p. 5). Logits is a way of measuring a point on a continuum of difficulty. (In adaptive tests the logit value represents a probability estimate that students will be successful at a certain ability level. That logit is converted to points for the recommended cutoff scores for the S-CAPE). BYU’s Spanish faculty then made a second check of items, and 51 additional items were removed from the test bank because they were seen as invalid test items.

The initial version of the S-CAPE began at Level 20 (of 51 levels) and increased 6 levels of difficulty if the student answered the question correctly or decreased 5 levels if the student got the question wrong. This is referred to as probing. After six questions, the levels increased or decreased by one in order to get a more precise measure of the examinee’s abilities, which was given to the examinee at the conclusion of the test. The test generally takes between 20–30 minutes to complete. Although the test may end in a few minutes if students answers five questions correctly or misses four questions at any level (VanBuren, 1994, p. 864). (Item difficulty is the only variable the S-CAPE uses to differentiate between levels, which hypothetically means that a student could get only grammar, vocabulary or reading comprehension questions without seeing any other types—a potential content validity concern.)

During the original validation process, cutoff scores were also created. A group of students who represented the ability levels of different classes served as a norming group from which standard scores could be applied. Once the cutoff scores were set, the S-CAPE began to be used as the primary means of placing students into Spanish classes for the BYU Department of Spanish and Portuguese.

In 1986, Larson (1991) used the Pearson Product Moment Correlation formula and calculated a reliability coefficient of .86. This was done after 43 students took the S-CAPE twice
on two consecutive days. That same year, 139 students took the S-CAPE at the beginning of Fall Semester. Midway through the semester, the instructors of these 139 students were asked to rate how appropriately these students were placed. Of those instructors, 79.9% stated that the S-CAPE had given a “good to excellent” placement with only three of those instructors reporting that the S-CAPE placed their students too high (p. 284).

The following figures were taken from the S-CAPE. The test begins with an introduction and a practice question as illustrated in Figures 2.1 and 2.2. These instructions are meant to teach students how to use the test to keep the test form from interfering with what students know.

![Brigham Young University Placement Exam](image)

**Figure 2.1. Introduction of the S-CAPE to Students.**
Students are also required to enter basic information such as their names, student identification numbers, and email addresses. It is assumed that the collection of this information was for the purpose of monitoring the S-CAPE. Students also have the option of seeing their results from the previous time they took the S-CAPE or of taking the test again.

The following are examples of questions from the S-CAPE that represent the testing of vocabulary, grammar, and reading comprehension (comprising Figures 2.3–2.8), although they appear in random order on the exam. Notice that all of the items on the test consist of multiple choice questions that require students to select the best answer from the context.
Figure 2.4. S-CAPE Vocabulary Sample Question 2.

Some of these questions are extremely basic, but there are questions spanning a difficulty level of 1 to 21. Figures 2.3 and 2.4 are both questions where the student needs to find a word that replaces the one used in the sentence. Vocabulary knowledge is also tested with questions where the student has to identify the object that best fits the description provided and by selecting the best translation of the idea of a given saying in Spanish.

Figures 2.5 and 2.6 present questions related to the testing of grammar knowledge.

Figure 2.5. S-CAPE Grammar Sample Question 1.
In the S-CAPE, grammar is typically tested with questions where the student fills in the blank with the best multiple choice option.

Last, the questions in Figures 2.7 and 2.8 test students’ reading comprehension. These questions take a longer time to complete because reading an accompanying paragraph is involved.
Some of the reading comprehension questions only require a brief scanning of the paragraph to identify the correct answer. Other questions require inferring to identify the correct answer. The topics used for the reading passages are based on general knowledge, so they can still be used despite the age of the test.

It is important to note that all of the items used for the S-CAPE assess content knowledge, not proficiency. Also, there are no questions that assess students’ listening or speaking proficiency.

Currently students only receive their score after finishing the S-CAPE, but a report, such as the one in Figure 2.9, illustrates how the S-CAPE works.
<table>
<thead>
<tr>
<th>Difficulty Level</th>
<th>Test Item</th>
<th>Student Answer</th>
<th>Right/Wrong</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>174</td>
<td>A</td>
<td>w</td>
</tr>
<tr>
<td>15</td>
<td>123</td>
<td>D</td>
<td>r</td>
</tr>
<tr>
<td>21</td>
<td>182</td>
<td>C</td>
<td>w</td>
</tr>
<tr>
<td>16</td>
<td>134</td>
<td>B</td>
<td>r</td>
</tr>
<tr>
<td>22</td>
<td>191</td>
<td>B</td>
<td>r</td>
</tr>
<tr>
<td>28</td>
<td>244</td>
<td>A</td>
<td>w</td>
</tr>
<tr>
<td>23</td>
<td>202</td>
<td>C</td>
<td>w</td>
</tr>
<tr>
<td>22</td>
<td>188</td>
<td>B</td>
<td>r</td>
</tr>
<tr>
<td>23</td>
<td>197</td>
<td>A</td>
<td>r</td>
</tr>
<tr>
<td>24</td>
<td>212</td>
<td>D</td>
<td>r</td>
</tr>
<tr>
<td>25</td>
<td>220</td>
<td>D</td>
<td>w</td>
</tr>
<tr>
<td>24</td>
<td>206</td>
<td>B</td>
<td>w</td>
</tr>
<tr>
<td>23</td>
<td>198</td>
<td>B</td>
<td>r</td>
</tr>
<tr>
<td>24</td>
<td>205</td>
<td>C</td>
<td>r</td>
</tr>
<tr>
<td>25</td>
<td>218</td>
<td>A</td>
<td>w</td>
</tr>
<tr>
<td>24</td>
<td>210</td>
<td>A</td>
<td>w</td>
</tr>
<tr>
<td>23</td>
<td>204</td>
<td>D</td>
<td>r</td>
</tr>
<tr>
<td>24</td>
<td>213</td>
<td>B</td>
<td>w</td>
</tr>
<tr>
<td>23</td>
<td>196</td>
<td>C</td>
<td>r</td>
</tr>
<tr>
<td>24</td>
<td>211</td>
<td>A</td>
<td>w</td>
</tr>
</tbody>
</table>

There were four incorrect answers at level 24
Number right: 10  Number wrong: 10
This student placed at level 23
Test completed at 12:07:52

Figure 2.9. Sample S-CAPE Administrator’s Report for an Individual Student (Larson, 1987, 23).

This type of report is only seen by an administrator. Students only see their score and the placement recommendation from that score. According to Larson (1986), the S-CAPE was set up with a flagging system to avoid students seeing the same question twice. If a question were used during the previous testing session, it would not be used during the subsequent testing session (p. 6).

Larson (1991) reported “initial observations and evaluations of the S-CAPE have been very positive” (p. 287). The following section looks at what others have said about the S-CAPE.
Responses to the S-CAPE

Generally, reactions to the S-CAPE have been positive. Burston and Monville-Burston (1995) stated that the S-CAPE was one of the best-known foreign language Computer Adaptive Tests (p. 42). VanBuren (1994) in her review of the S-CAPE, also reacted favorably and even recommended the exam for teachers to evaluate for their schools. However, she also stated that the exam did put too much emphasis on literary vocabulary and that it lacked a listening, speaking, and writing evaluation. Additionally she points out that some of the questions lack cultural relevance. Her recommendation for future versions was that the vocabulary reflect more of “the daily activities of native Spanish speakers and [be] drawn from a broad base of authentic sources other than literature” (p. 865). However, Lam (2010) stated that although the S-CAPE is standardized and validated, it may not align with the objectives of another university’s program, and it is limited in its scope.

Testing for Intermediate and Advanced Proficiency Students

There is minimal research on the placement of advanced students. What has been done either focuses on heritage speakers, or needs to be updated. Intermediate levels of language proficiency can be difficult to assess. The S-CAPE only claims to be able to place from novice to third semester language students (Perpetual Technology Group, 2008). Lam (2010) concluded that it may be impossible to place intermediate level students due to the fact that gains in language learning may not be as observable at this level. This may also be due to the type of testing that has been available in language learning. After developing the R-CAPE, a CAPE test for Russian, Larson stated, “Until answer-judging routines via artificial intelligence become more refined, we are basically limited to testing receptive skills only, meaning that acceptable computerized tests of speaking and writing are not yet possible” (Larson, 2000, p. 52). However,
advancements in the field of language testing show that these could be possibilities in the future. For example, Eggington and Cox (2013) showed the Elicited Oral Response tests (EOR) along with Automatic Speech Recognition (ASR) are a valid, reliable and practical way of assessing speaking skills. “The evidence suggests that ASR-scored EOR tests could be used to predict speaking ability, especially in making decisions such as placement testing. [The evidence] also seems to show great potential as a cost-effective alternative to conducting expensive face-to-face speaking-proficiency interviews” (p. 137). These, and other advancements, could increase the effectiveness of placement testing.

**Research Questions**

Placement tests, as other assessments, need to be reliable, valid, and meet the needs of the context in which they are used. For that purpose, the following questions were used to assess the S-CAPE within the context in which it is being administered.

1. To what extent does the S-CAPE still function as a reliable assessment instrument?
2. How effective is the S-CAPE for placing students into SPAN 206 and 321 at Brigham Young University?
3. What factors affect the functionality of the S-CAPE in making correct placement decisions?

**Chapter Summary**

Placement tests are a valuable means of placing language students. The S-CAPE was developed and found reliable and valid for this purpose, and it has been well received by some groups and sold to many others. However, since it was not originally developed to place intermediate and advanced students, it needs to be tested to see if it is capable of such placement,
and if such placement is even possible. Future developments in the field will enable placement testing to include written and oral assessments in more feasible forms.
CHAPTER 3

Research Design and Methods

The purpose of this study was to examine the reliability of the S-CAPE, its effectiveness in placing students in upper level Spanish courses, and to explore what other factors could affect the functionality of the exam. To answer these questions, three primary sources of data were collected: students’ results on the S-CAPE, professors’ and instructors’ reactions to how their students performed on the S-CAPE, and the results of a student survey.

Methodology

This research study was conducted during Fall Semester of 2013 at Brigham Young University. To answer the first research question about the exam’s reliability, participants took the S-CAPE twice for a test-retest analysis. To make the test experience similar for all participants, the class syllabi for SPAN 105, 106, 205, 206, and 321 notified students that one of their class assignments would be to take the S-CAPE for the purpose of updating the placement test (see Appendix A for information included in the syllabi). The class syllabi also specified that in order for students to receive credit, they would be required to take the exam twice. However, even though students were required to take the S-CAPE, they had the option to not have their results included in this study, and a few students chose that option. The school semester began September 3, 2013, and the test was scheduled for September 16–27. It was purposefully prepared to begin after the add/drop deadline due to high numbers of class changes before this time. The end date was then extended until October 4 per instructor request, as well as to ensure that as many students participated as possible. Student ID numbers were used in order to connect the first time students took the test to the second time they took the test, as well as to ensure that students received credit for test completion. To further ensure the reliability of the results,
students were required to take the test in the Humanities Testing Lab (HTRSC) located in the basement of the Joseph Fielding Smith Building on the BYU campus. This was done to increase the reliability of the test results by ensuring that students were identified and taking the test in a secured environment.

It was assumed that students were actually trying their best as they took the S-CAPE exam; however, students were told that they would receive credit for taking the S-CAPE twice and not for the score that they received.

To evaluate the effectiveness of the S-CAPE in placing students, instructors were surveyed after all students had taken the S-CAPE. This was to assess the face validity of the S-CAPE, or in other words, how instructors felt about the S-CAPE as an assessment tool. When the S-CAPE data was received, each teacher received his or her students’ S-CAPE results. This was done to give credit to students who had taken the S-CAPE, but more importantly, a survey was sent to teachers asking them to decide how many students the S-CAPE had placed correctly. To answer this question, and along with the students’ scores that were provided, instructors were given the current recommended cutoff scores used by the BYU Department of Spanish and Portuguese.

All instructors whose students participated in the S-CAPE study were sent the survey via email and responses were received the same way. Consolidated data did not include any information that identified participants. The survey also asked instructors for additional feedback about the S-CAPE.

To further explore the effectiveness of the test and other factors that could affect the functionality of the test in correct placement of students, students were also surveyed about their
placement. There were two student surveys. The first survey was taken before the S-CAPE and
the second survey was administered to students at the end of the Fall Semester.

Participants

Participants included students at BYU who had enrolled in Spanish classes ranging from
SPAN 105–321 for Fall Semester 2013. They also included the instructors that taught those
courses. Roughly 1,000–1,500 students took these courses during that semester. Of those, 727
students participated in the S-CAPE study, allowed their results to be included in this research,
and took the S-CAPE twice. These students were between the ages of 17–29 (mean = 21) with a
few students in their 30s and one in her 50s. For the S-CAPE post-survey that was distributed
after the semester, a total of 311 students responded. Table 2 shows the distribution of how many
students were originally registered for each course, disaggregated by gender. Out of the 321
student responses, the only comments that are included in this thesis come from students who
originally agreed to participate in the S-CAPE study at the time they took the S-CAPE.

Table 2

Demographics of Students Participating in the S-CAPE Study

Class registration, disaggregated by gender

<table>
<thead>
<tr>
<th></th>
<th>SPAN 105</th>
<th>SPAN 106</th>
<th>SPAN 205</th>
<th>SPAN 206</th>
<th>SPAN 321</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>63</td>
<td>31</td>
<td>54</td>
<td>30</td>
<td>91</td>
<td>269</td>
</tr>
<tr>
<td>Male</td>
<td>23</td>
<td>4</td>
<td>10</td>
<td>9</td>
<td>412</td>
<td>458</td>
</tr>
<tr>
<td>Total</td>
<td>86</td>
<td>35</td>
<td>64</td>
<td>39</td>
<td>503</td>
<td>727</td>
</tr>
</tbody>
</table>

As the table demonstrates, the majority of students enrolled in each course were female, with the
exception of SPAN 321, where the trend was completely reversed. The higher enrollment in the
upper level classes is most likely due to more returned missionaries being male. This has been the norm in previous years, but the number of women who are serving missions is increasing.

To evaluate the effectiveness of the placement decisions, instructors were surveyed. There were 22 who were sent the survey. Of the instructors recruited, 8 responded. Table 3 identifies their demographic.

Table 3

Demographics of Instructors Participating in the S-CAPE Study

<table>
<thead>
<tr>
<th>Class taught, disaggregated by gender</th>
<th>SPAN 105</th>
<th>SPAN 106</th>
<th>SPAN 205</th>
<th>SPAN 206</th>
<th>SPAN 321</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Male</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

Half of the instructors represented SPAN 321. Also, the instructors for SPAN 105–SPAN206 were all graduate students at BYU.

Data Instruments

As mentioned previously, data was collected from three different sources: the S-CAPE, the student surveys, and the instructor surveys. As was also previously mentioned, the S-CAPE is a computer-adaptive test of grammar, reading, and vocabulary used to advise students regarding the class in which they should enroll. Interested readers should refer to the detailed description of the S-CAPE that has already been discussed in Chapter 2. The S-CAPE was administered twice for test-retest reliability data.

Two student surveys were administered to generate data. The first, a pre-S-CAPE survey, was administered twice—once each time prior to the test-retest S-CAPE. This survey inquired
about factors that could influence students’ proficiency, including participants’ age, gender, current Spanish class, number of years studied, and the way that they learned Spanish (see Appendix B). Students were encouraged to take the survey before they took the S-CAPE, but they did not necessarily need to complete the survey to take the test. It was assumed that if students answered “yes” to a survey question and then left it blank the second time they took the test, that the information had not changed, but that students had opted to not fill in the answers the second time around. There was no difference between the content of the survey the first and the second time that students took the exam.

After receiving the pre-survey results, it was discovered that a computer error had erased half of the survey data for every student. No S-CAPE data was lost due to the computer error. Because of this data loss, a second survey was created that added additional questions. This survey was created and administered using Google Forms. The questions used are in Appendix D.

This survey was sent out during the break between Fall Semester 2013 and Winter Semester 2014. Students were initially contacted during the beginning of the break, and then reminded to complete the survey at the end of the break. It used the questions asked in the pre-S-CAPE survey with the addition of questions that asked more about returned missionaries’ language use on their missions and how students felt about the S-CAPE as a placement tool. All of the data received was analyzed.

**Data Analysis**

To answer the research questions, three different types of data analyses were used for the S-CAPE data. Each phase fulfilled a different purpose: a test-retest analysis was used to determine reliability, followed by an analysis to determine the extent to which the test accurately
discriminated among different proficiency levels. The analysis concluded with an examination of
the degree to which teachers felt that students’ placement results adequately reflected their
performance. Each of these phases of analysis will be described in more detail in the sections
below.

During the first phase of analysis, the reliability of the S-CAPE was measured. This was
measured using test-retest reliability. Only students who took the S-CAPE twice were included
as two test scores were needed in order to perform this analysis. When there were more than two
scores, only the first two were used. The two S-CAPE scores were analyzed based on the idea of

To answer the second question regarding the effectiveness of the S-CAPE to differentiate
student proficiency at higher levels, an ANOVA was conducted and effect size was evaluated
with Cohen’s $d$. If the test score mean of students is higher than the mean of the students in lower
levels, then the S-CAPE might differentiate at the 300 level well enough to be used as a
placement tool. For example, many amusement parks have minimum height requirements for
riders. Potential riders being measured twice with a carpenter’s ruler might have different results
on each measurement, but if the outcome still permits those tall enough to get on and those who
are not to wait until they reach the required height, the discrimination based on the cut score
would be sufficient (Cox, 2014, personal communication). If such a difference exists, then this
placement test could reasonably be used for returned missionaries. If not, then further research
could determine the possibility that a test with more items at the advanced levels could increase the accuracy of placing higher level students. It may also be that neither this nor any discrete point test is capable of placing students at the intermediate and advanced levels.

The third question asks about additional factors that affect the functionality of the S-CAPE. Surveys were used as a way to assess how students and instructors felt about the face validity of the S-CAPE. They were also used to pinpoint other factors that would affect the S-CAPE’s functionality. The pre-survey was used primarily to connect students and their S-CAPE scores to the class in which they had enrolled. The post-survey was analyzed to see how students chose their Spanish class, if students’ comments expressed that the S-CAPE was effective or not, how many students followed the S-CAPE’s results, and if they felt that it placed them correctly. These analyses were also organized by the class in which the students were enrolled. The placement options provided to students on the survey were S-CAPE, It was the next course in the sequence, By its course description, Teacher/professor recommendation, and Other. If students chose Other, then results were categorized to find placement patterns. For the teacher survey, results were separated by the level the instructor taught for analysis.

In each of these analyses, students’ names and ID numbers were used to connect students’ data. After that, any names associated with the survey results, including students’ individual comments, were eliminated to preserve student anonymity.

Chapter Summary

The methodology and the data analysis of this S-CAPE study were set up to enable the BYU Department of Spanish and Portuguese to draw conclusions about the use of the S-CAPE in placement decisions. If the S-CAPE is a valid and a reliable test, it will also be tested to see if it is capable of placing students with intermediate or advanced proficiency.
CHAPTER 4

S-CAPE Findings

In this chapter the findings of the S-CAPE study are presented. In order to be a valid test for the current BYU context, the S-CAPE needs to differentiate between class levels and the teachers of the Spanish courses need to believe that students’ results are justified. Students also should feel the test recommendations are well founded. Survey results revealed how both teachers and students felt about the validity and reliability of the S-CAPE. Once again, the focus was on the following research questions, and the information presented was organized around them:

1. To what extent does the S-CAPE still function as a reliable assessment instrument?
2. How effective is the S-CAPE for placing students into SPAN 206 and 321 at Brigham Young University?
3. What factors affect the functionality of the S-CAPE in making correct placement decisions?

The results may or may not be generalizable to other placement contexts. There were 726 S-CAPE students’ results that were included in this study. Currently, 520 is the recommended cutoff score for placement into SPAN 321. The mean S-CAPE score for all students was 596.91 (SD = 142.08, 95% CI [586.49, 607.33]). This finding is not surprising as 69% of the subjects (see Table 3) were enrolled in SPAN 321. Figure 4.1 shows the distribution of the mean scores from the S-CAPE divided by 100-, 200- and 300-level courses.
Figure 4.1. Distribution of the Mean of S-CAPE Scores Divided by 100-, 200- and 300-Level Courses.

This figure represents the mean results of 727 students. As mentioned previously, SPAN 321 had the most S-CAPE results because that course had the largest number of students enrolled in it.

The following section presents the reliability of the S-CAPE as determined by the Pearson product-moment correlation.

Reliability of the S-CAPE

The test-retest method was used here to re-examine the reliability of the S-CAPE. Each participant’s results that were included in this study had two test scores. The Pearson product-moment correlation for students’ results was .64 (p < .01). [Correlation is significant at the 0.01 level (2-tailed)]. Originally, the author of the S-CAPE found the reliability to be .86 (Perpetual Technology Group, 2008). Part of the difference between reliability results could be explained by the differences in the two studies. For example, Dr. Larson (1991) had a study of 43
participants who were most likely evenly divided between the courses for which the S-CAPE was designed. For this current study, over 727 participants’ results were included and a little more than half of those were from SPAN 321, a course which was not originally part of the design of the S-CAPE. Figure 4.2 presents the correlation between the first and second test score of the S-CAPE participants.

![Figure 4.2. Results of the Pearson Product-Movement Correlation.](image)

This figure shows that the most common result was there were little to no difference between the two scores. This is illustrated by the fact that most of the results fall on or close to the middle line in the figure. This result supports the validity of the S-CAPE and the idea that students were actually trying and not just rushing through the test. However, this figure also shows variation
with the greatest difference in variability over the 600 point range, as illustrated by the dotted line. This range of skills is most representative of scores from students with skills over the SPAN 321 range. In other words, the reliability of the S-CAPE decreases at the SPAN 321 level.

Students with such varying results were confused at the differences in their scores. Representative comments on students’ surveys included: “It gave me wildly different results;” “The results were very different from the two times that I took it and I don’t think it would have accurately placed me because if I had taken a more advanced class I would not have been prepared for the grammar;” “I found it very helpful once I took it multiple times. I averaged my score out of the 3 times I took it. If I had only taken it once, I do not think it would have been very accurate;” and “It is not always very accurate. At one time I was placed in the 105 class, and another time I was placed in 205. I feel that some of the questions were poorly worded so that I did not understand what it was asking.” These students’ comments, and many others, showed that students did not feel that their test scores were reliable.

As mentioned previously, it is expected that placement tests have reliability greater than .80 (Carr, 2011). The reliability of the S-CAPE was .64, which means that over one third of the variance in test scores was due to measurement error. This could be due to the design of the adaptive test. As reading, vocabulary, and grammar are all treated as a unidimensional trait with items being selected solely based on item difficulty, the algorithm may have given a student with mostly grammar items one time and mostly vocabulary another. The unidimensional treatment of the skills may have affected content validity. Students would need to take the test four times and another test-retest interrater reliability would need to be run before students could be told that by taking the S-CAPE two times and averaging their scores they would have a more accurate indication of what class they should be in (Eggett, 2014, personal communication).
**The S-CAPE and Placement**

Levene’s test was used to look at the homogeneity of variances of the placement data from the various Spanish classes involved in this study. An ANOVA test was used to compare the mean scores from each course and to see if there was statistical significance between the different classes, and was followed by a Bonferroni post-hoc test to analyze how well the S-CAPE was able to place students into each course. Cohen’s $d$ of the effect size was also reported. However, first the confidence interval was reported. Figure 4.3 is a representation of the confidence interval based on a 95% confidence level.

![Figure 4.3. 95% Confidence Interval Based on the Mean and Standard Deviation of Students in SPAN 105–321.](image)

Notice the similarities between SPAN 105-206 and the difference at the SPAN 321 level.
Levene’s test for homogeneity of variances of the Spanish courses’ placement data \((F = 23.72, p < .001, df = 4, 722)\) indicated that equal variances could not be assumed. Upon further investigation, it was determined that there was large variance with the large sample size. Consequently, so the potential negative impact was negligible, so an ANOVA with its subsequent post-hoc Bonferroni contrast was conducted as well as an analysis of effect which has fewer constraints than ANOVAs.

Table 4 presents ANOVA analyses for SPAN 105, 106, 205, 206, and 321.

Table 4

**ANOVA: Single Factor Summary for SPAN 105–321.**

<table>
<thead>
<tr>
<th>Groups</th>
<th>Number of Students</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPAN 105</td>
<td>86</td>
<td>422.02</td>
<td>50.27</td>
</tr>
<tr>
<td>SPAN 106</td>
<td>35</td>
<td>432.99</td>
<td>58.77</td>
</tr>
<tr>
<td>SPAN 205</td>
<td>64</td>
<td>480.77</td>
<td>94.79</td>
</tr>
<tr>
<td>SPAN 206</td>
<td>39</td>
<td>497.82</td>
<td>77.25</td>
</tr>
<tr>
<td>SPAN 321</td>
<td>503</td>
<td>660.67</td>
<td>115.25</td>
</tr>
</tbody>
</table>

ANOVA

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>(F)</th>
<th>(P)-value</th>
<th>(F) crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>6862083.50</td>
<td>4.00</td>
<td>1715520.87</td>
<td>158.93</td>
<td>0.00</td>
<td>2.38</td>
</tr>
<tr>
<td>Within Groups</td>
<td>7793472.80</td>
<td>722.00</td>
<td>10794.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14655556.30</td>
<td>726.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is interesting that the mean scores for SPAN 105–206 are all in the 400 point range. This is interesting because there is not a lot of difference between these four classes on the difficulty continuum of the S-CAPE. However, the mean of the classes also shows that there is some standard distance between classes. Additionally, as classes become more difficult, so does the variability in the class, with the greatest variability being in SPAN 321 (except for SPAN 206, which shows less variability than 205). This is not too surprising as the test does not include
an oral component, and thus there would be greater variability among returned missionaries who might speak well but struggle with literacy. There was a statistically significant difference between classes as determined by one-way ANOVA ($F(4,722) = 158.93, p = 0$). The $p$-value of 0 shows that there was a lot of variation between scores, and therefore justifies the use of a post-hoc test (t-test) to measure that variability between classes.

Comparisons using Bonferroni’s contrasts found no statistical difference between the SPAN 105 and SPAN 106 students (mean difference = -10.96, a 95% CI between -69.62 and 47.69, and $p = 1.00$) with Cohen’s $d$ indicating little to no effect ($d = .2$).

Results were similar for SPAN 106 and 205. While it is evident from Figure 4.3 that the difference between the 106 and 205 students is larger than previous pairing, it is still not significantly different (mean difference = -47.78, a 95% CI between -109.28 and 13.72, and $p = .29$). However, the effect size between the two groups was large with Cohen’s $d = .61$.

Similar to the 100 level classes, the difference between SPAN 205 and SPAN 206 was not significant (mean difference = -17.05, a 95% CI between -76.48 and 42.37, and $p = 1.00$) with Cohen’s $d$ indicating little to no effect ($d = .2$).

Results were different for SPAN 206 and SPAN 321. Results showed that there was a statistically significant difference between the two courses (mean difference = 162.85, a 95% CI between 114.23 and 211.48, and $p < .001$) with Cohen’s $d$ indicating a large effect size ($d = 1.66$).

While the S-CAPE is good at differenchiating between whether or not students should be placed in the 100 level as opposed to the 300 level, it is not good at discriminating in more fine-grained ways (i.e., whether students should be placed in SPAN 105 or 106.) Therefore, if the reliability of the S-CAPE was increased, it could be capable of placing into SPAN 321, but it
may not be good at placing at the levels it is currently placing at. Conclusions about the S-CAPE being capable of placing beyond SPAN 321 are beyond the scope of this research.

**Analysis of Perceived S-CAPE Validity**

The analysis of the validity of the S-CAPE began with the results of the survey completed by class instructors and then continued with the results of the survey completed by student participants. This helped measure the face validity of the S-CAPE, or in other words, the extent to which the exam appeared to accurately measure what it claims to measure.

**Survey of the correct placement of the S-CAPE according to class instructors.** To assess the face validity of the S-CAPE from a teachers’ perspective, the instructors of SPAN 105–321 were surveyed. They were provided with students’ scores as well as the current placement recommendations used by the BYU Department of Spanish and Portuguese. It was assumed that how instructors felt about the S-CAPE’s ability to place students would affect the extent to which they recommended it or use it as part of their course. Of the 22 instructors that were contacted by email, eight responded, which represented 36% of instructors. This was a low rate of return, and survey results may not be generalizable to other instructors. The results from those eight instructors showed that they felt that in total, 86% of their students were placed correctly in this study. This was higher than the researcher had anticipated because before this study was conducted, many students and instructors gave the researcher negative feedback about the S-CAPE. If the instructors’ results are disaggregated further, we see that the instructors who responded from SPAN 105–206 felt that the S-CAPE placed an average of 80% of their students accurately, while the instructors from SPAN 321 felt that an average of 92% of their students were placed accurately. Those who taught higher-level classes expressed more confidence in the
S-CAPE’s ability to place their students than the teachers of lower-level classes. Instructors were also asked if they had any further feedback about the S-CAPE, but none commented.

Survey of the effectiveness of the S-CAPE according to students. In addition to asking instructors for their feedback, students were also surveyed, and their responses are presented in this section. A survey was sent out at the end of Fall Semester 2013 to all students taking SPAN 105–321. Participation was voluntary, and there were 311 students who responded. Five students responded twice, but their results and comments about the S-CAPE were consolidated into one survey to prevent their opinions from being counted twice.

In response to the survey question “Did you take the S-CAPE at the beginning of Fall Semester 2013?” 189 students out of 311 responded affirmatively. Of those, 135 (71.4%) reported that they followed the S-CAPE’s results. Another 152 (80.4%) students reported that they felt like it placed them correctly. These results are confusing because the number of students that reported that the S-CAPE placed them correctly was higher than the number of students that reported that they followed the S-CAPE’s results. Students may have been confused by the wording, and they may have included any time they took the S-CAPE during the assigned semester as part of their results. However, the results still suggest a more positive attitude towards using the S-CAPE than was anticipated.

At the end of the survey, students were able to answer, in an open-ended format, the following question: “How do you feel about the S-CAPE as a placement tool?” After looking at the responses, the author decided that students were expressing the following ideas about the S-CAPE: Effective, Has Effective and Ineffective Parts, Neutral, Ineffective, and No Comment. (The category of No Comment also includes a few rare cases of students who never actually took the S-CAPE, but who opted to take the survey.) Comments not included in this section are those
that dealt with students’ reactions to the research study (some students either did not understand the purpose of the study or did not understand why it took place when it did during the semester.) As demonstrated by Figure 4.3, almost half of students indicated that they felt that the S-CAPE was an effective placement tool.

![Out of 311 survey responses](image)

**Figure 4.4. How Students Felt About the Effectiveness of the S-CAPE as a Placement Tool.**

Once again, the results surprised the researcher because they showed that students had more confidence in the S-CAPE than was previously expected. In terms of percentages, and combining *Effective* and *Has Effective and Ineffective Parts*, 59% of students felt that the S-CAPE placed them with some degree of effectiveness. Adding *No Comment* then brings the percentage up to 80%. Although students questioned the S-CAPE’s reliability, because their results were so varied from one testing to another, they still seemed to feel that the S-CAPE did place them correctly.

**Factors that Changed the Functionality of the S-CAPE**

The purpose of the subsequent section is to analyze how the S-CAPE fits into the placement process. As in the previous section, students’ survey results were analyzed to see how
they placed themselves into the class they were taking at the time of the S-CAPE study. It also looks at the S-CAPE results of returned missionaries to see how they compared to non-returned missionaries at the SPAN 321 level in order to assess the difference between the two groups.

**How students placed themselves.** One goal of the survey sent out to students was to determine how they decided which Spanish class to take. Students could indicate one of the following placement options on the survey: S-CAPE, *By its course description*, *It was the next course in the sequence*, *Teacher/professor recommendation*, or *Other*. Only 4 students chose S-CAPE, where 138 chose *By its course description*, 93 selected *It was the next course in the sequence*, 23 indicated *Teacher/professor recommendation*, and 53 selected *Other*. The following table illustrates students’ responses:

![Figure 4.5. How Students Chose Their Current Class; Survey Options.](image)

Almost half of the students surveyed indicated that they chose their class by its course description, whereas the S-CAPE accounted for only 1% of the responses. This does not mean that students did not take the S-CAPE, but rather that they did not select it as their primary means of placement. In the *Other* category, four students indicated that they made their selection based partially on the S-CAPE. The only students to comment on using the S-CAPE were those
fromSPAN205and206. The SPAN 321 students had the most Other comments, meaning that they felt like they placed themselves in other ways than those that were provided for them in the list of options; however, there were also more students that took SPAN 321 than any other class. Students of SPAN 105 mostly placed themselves by the course description, and then the course after (SPAN 106) had participants that principally relied on the fact that it was the next course in the sequence. (SPAN 105 and SPAN 205 tend to be entry points for students and then they continue through the course sequence afterward.)

The 53 students who selected Other were able to write an explanation for their answers. After results were received, the researcher organized these explanations in the following manner: Partially based on the S-CAPE, Friends, Family, Word of Mouth/Returned Missionary Class, 16-Credit Exam, Transfer Student, Counselor Recommendation or Wanted to Learn. The Returned Missionary Class is the idea that students connote SPAN 321 as being the class that returned missionaries automatically take. The 16-Credit Exam category means that students enrolled in the Spanish class in order to take the 16-Credit Challenge Exam. This exam tests Spanish knowledge, which can give students college credit for Spanish learned elsewhere. One of the department regulations is that students be registered for a Spanish class and that it be their first Spanish class on campus. It also needs to be a class at the 200-level or higher in order to qualify to take this exam and to receive credit. (If students are registered for a class below SPAN 321 it affects how many credits they can receive from the exam). Figure 4.4 shows the number of student responses in each category.
Figure 4.6. Results of the Other Category.

One vote was included in both the Family and the Friends category due to the fact that it had both in its description. Reliance on non-professional opinions made up 58% of the Other category; Friends (12 responses), Word of Mouth/Returned Missionary Class (11 responses), and Family (8 responses). Not included in Figure 4.4 were the following responses that were left uncategorized: “I didn’t want to take Spanish 101 or 102;” “I’m a native speaker and I need that class for my major;” “I looked up which course I should take online;” “Personal workload and difficulty preference;” and “Status quo.”

Limitations of This Study

There are several limitations to this study in terms of students, instructors and the researcher. Students were organized based on the classes in which they placed themselves, which means that their ability level could be due to more than the S-CAPE results. In addition, the study was based on the assumption that students were actually trying to do well on their two attempts of the S-CAPE. It should also be noted that the number of students who participated in this study from SPAN 321 was a little more than double all of the students who participated from SPAN 105–206 combined. This may have affected the research and could explain some of the discrepancy between the reliability of the S-CAPE that Dr. Larson found, which was .86 (1991)
and the reliability found in this study, which was .64. Furthermore, the feedback given by instructors was helpful to this study, but it should be taken into account they were mostly student instructors and adjunct faculty with a few professors who may not have had the professional training to judge the validity of a placement exam. Specifically for the student instructors, they may be making a lot of assumptions since they are employed by BYU and commenting on a test created at BYU. The bias of the researcher at the beginning of this study was that the S-CAPE, based on feedback from faculty as well as students, was not placing students as it was designed to do, and that it needed to be updated. However, the researcher controlled for this bias by not interfering with the data that was collected.

Chapter Summary

In this chapter, the reliability and the validity of the S-CAPE were presented. It was found that the Pearson product-moment correlation for the S-CAPE was lower than expected for the new student population ($r = .64$). While many students and teachers felt the test was effective, some students also questioned the reliability of the S-CAPE and felt they needed to take it multiple times. However, the results of the S-CAPE did indicate that students in the SPAN 206 and 321 levels had significantly different test scores, indicating that it might function well enough in making general upper level distinctions if the reliability could be improved. In terms of validity, class instructors and students showed confidence in the face validity of the S-CAPE. However, there were other factors that seemed to be more dominating in how students actually ended up placing themselves, which were friends, family, and the course sequence.
CHAPTER 5

Discussion and Implications

In Chapter 4 of this thesis, the data from students’ S-CAPE results and surveys were analyzed, as well as how instructors felt that the S-CAPE placed their students. The following section interprets that data and looks at the implications for the Department of Spanish and Portuguese at BYU. The first section of this chapter comments on the reliability of the S-CAPE, on the confidence that teachers and students had in the S-CAPE, and how students actually placed themselves. Recommendations to improve the S-CAPE are made. The pedagogical implications provide suggestions to improve the overall placement system in BYU’s Department of Spanish and Portuguese.

The Reliability of the S-CAPE

As presented in Chapter 4, the reliability of the S-CAPE was low (.64). Students can actually be more confused about which Spanish class to enroll in after taking the S-CAPE than before taking it. This is exactly what the BYU Department of Spanish and Portuguese is trying to avoid. The S-CAPE’s reliability also affects its ability to place intermediate and advanced level students. However, it cannot be concluded that the S-CAPE is capable of placing at these levels until the S-CAPE’s reliability can be increased. (See “Suggestions for Changes to the S-CAPE” in this chapter.)

Teacher and Student Confidence in the S-CAPE

Only 36% of teachers responded to the survey about how well they felt the S-CAPE placed their students. However, it is interesting to note that instructors’ and professors’ confidence in the S-CAPE increased the higher the level of the class they taught. The author presents two possible explanations for this. First of all, due to the nature of language
development and placement, it may be easier to identify the errors in the placement system at a lower level. For example, if a student who has a very minimal knowledge of Spanish is placed into SPAN 106, it is obvious that he/she has been placed into the wrong class. Whereas a student in SPAN 321 who really belongs in SPAN 205 may not stand out as much when compared to his/her classmates. Thus, professors and instructors could feel that the placement test was correctly placing students where in reality a misplaced student’s lower skill level is not as noticeable in more advanced courses.

The second possible explanation is connected with the cutoff scores. Currently, SPAN 321 students only need to obtain 520 (converted from a logit score which estimates the success of students on a difficulty continuum) for the test to recommend that they be placed into that same class. Students in SPAN 321 scored between 520 and 950. This variance in SPAN 321 is greater than other classes. Thus, at the SPAN 321 level it is easier to feel that the S-CAPE placed students correctly because of the wide variance allowed for that class.

Despite these issues, the instructors’ and professors’ confidence in the S-CAPE was much higher than expected. Nonetheless, it would be interesting to see if these same instructors were actually recommending the S-CAPE to their students.

Students also showed more confidence in the S-CAPE than expected. Of students surveyed, 152 out of 311 said that they followed the recommendations given to them by the S-CAPE for Fall Semester 2013 and that they felt that the test placed them correctly. If the 311 students are a representative sample of students taking Spanish classes, and the question was correctly interpreted by students, then more than half of BYU students are taking the S-CAPE to help them in their placement decisions. However, results were lower when students responded to the question about how they chose the class they were taking. Only 1% of students said that they
based their decision on the S-CAPE. Instead, they are relying on course descriptions and recommendations from others. A possible explanation for these results is the fact that students may have used the S-CAPE as one of multiple sources that went into their placement decision, but they may have felt that it was not the main source. For this latter question, students were only allowed to select one answer, but were able to explain an answer if they chose Other. If they had been allowed to select more than one option, perhaps the S-CAPE would have been included more. However, if students are not using the S-CAPE to make decision, what resources are they using?

**How Students Place Themselves**

Students of certain classes tended to follow a certain pattern in placement. Figure 5.1 shows how students placed themselves divided by class.

![Figure 5.1. How Students in SPAN 105–321 Placed Themselves.](image)

Notice that students of SPAN 105 mostly placed themselves by the course description, and students of SPAN 106 principally relied on the fact that it was the next course in the sequence. Hence, as soon as students are in the system, they generally follow the course sequence. (The author has seen some exceptions of students skipping courses or retaking classes.)
The author taught SPAN 105 here at BYU as a student instructor for three semesters and then began teaching SPAN 205. She was expecting her SPAN 205 students to be more advanced than her SPAN 105 students, but in many ways they were not. One of the reasons for this is that SPAN 205 can be an entering point for high school students, meaning they have had a previous background in Spanish, but they are not always prepared for the rigors of a university class. Changes to the overall placement system need to address these placement difficulties.

Figure 5.1 also shows that student in SPAN 321 had the most Other comments, meaning that they tended to use alternative sources to place themselves over those provided by the university. (Conversely, they also had the greatest number of students, meaning that statistically there would be more variation.) Many students followed the advice of friends and family members, which is not necessarily bad, but it keeps the Spanish and Portuguese Department at BYU from being able to influence the placement process. The author recommends the department assess all the resources that students mentioned that they used to make placement decision and ensure that accurate placement information is being distributed.

There are not a lot of resources available to students at the SPAN 321 level. The S-CAPE has not been typically recommended at this level by faculty and staff. This means that the placement resources available to students who consider taking SPAN 321 are the information off of the course registration, university counselors, instructors of the course, or sources outside of the university. This study obviates the need for placement resources for the SPAN 321 level. The S-CAPE could act as a placement tool if updates were made. Also, as will be discussed later in this chapter, self-placement tools may also be a good placement resource at this level.

Forty-four percent of students said that they chose their course by its course description. Figures 5.2 and 5.3 are two of the course descriptions off the BYU Route Y MyMAP: Academic
Planning website (March 24, 2014) for both SPAN 205 and SPAN 321. This is a sampling of what information students are given to choose their classes.

### SPAN 205 - University Spanish

<table>
<thead>
<tr>
<th>Description</th>
<th>Continued development of grammar, cultural understanding, conversation skills, writing, and reading through the study of literature. Students are expected to attain an Intermediate-High proficiency level.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credit Hours</td>
<td>4.0 credit hours, 5 class hours a week, and 1 lab hour</td>
</tr>
<tr>
<td>Headers</td>
<td>SPAN 205, all sections, can be taken by students who completed SPAN 201 W, Sp, Su 2007. SPAN 205 is for students with four years of high school Spanish or who have completed SPAN 106 or the equivalent.</td>
</tr>
<tr>
<td>Prerequisites</td>
<td>Four years of high school Spanish, SPAN 106, or placement by test. Four years of high school Spanish, SPAN 106, or placement by test.</td>
</tr>
<tr>
<td>Recommended</td>
<td>None</td>
</tr>
<tr>
<td>Note</td>
<td>Fulfills the University Core Languages of Learning requirement. This course is part of a GE Mosaic. See <a href="http://ge.byu.edu/mosaic-list">ge.byu.edu/mosaic-list</a> for more information.</td>
</tr>
<tr>
<td>When Taught</td>
<td>Fall; Winter; Spring; Summer.</td>
</tr>
</tbody>
</table>

**Figure 5.2. Course Information for SPAN 205.**

In the “Prerequisites” section, the placement test is mentioned, but it also assumes that students will know that this is referring to the S-CAPE and how they can access this test. This information also gives students the impression that the number of years they studied in high school or the previous classes taken are more important than the S-CAPE results. The description is rather
vague and it could include a hyperlink with the reference to the ACTFL (American Council on the Teaching of Foreign Languages) proficiency guidelines due to the fact that it mentions an Intermediate-High proficiency level with no other referent for students who may not be familiar with ACTFL. Figure 5.3 shows how SPAN 321 is presented to students who are registering for a Spanish class.

<table>
<thead>
<tr>
<th>SPAN 321 - Third-Year Spanish Reading, Grammar, And Culture.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td><strong>Credit Hours</strong></td>
</tr>
<tr>
<td><strong>Headers</strong></td>
</tr>
<tr>
<td><strong>Prerequisites</strong></td>
</tr>
<tr>
<td><strong>Recommended</strong></td>
</tr>
<tr>
<td><strong>Note</strong></td>
</tr>
<tr>
<td><strong>When Taught</strong></td>
</tr>
</tbody>
</table>

**Figure 5.3. Course Information for SPAN 321.**

Under “Prerequisites” in this section, the placement test is mentioned; however, it is not actually a prerequisite to take this course. This misinformation also may explain why more students did not state that they relied on the S-CAPE as a placement tool; they may not have known about it or may have trusted another’s opinion instead of taking the time to take it. Often, university counselors are blamed for recommending that returned missionaries enroll in SPAN 321 even
though they might not be ready for it, but survey results show that students are more inclined to use the information presented on the registration page (Figure 5.3) than other sources. Therefore, it is necessary to update this registration information, including the phrase “First class for returned Spanish-speaking missionaries.” SPAN 321 no longer meets the needs of all returned Spanish-speaking missionaries due to the number that serve within the States. However, because of the way that missionaries are taught language, they have had practice in analyzing their own language abilities, and therefore would benefit from having a list of statements that show what can be done at each level (see recommendations for Can Do Statements on Page 61).

In addition to improving the S-CAPE, changes can also be made to the overall placement system at BYU.

**Suggestions for Changes to the S-CAPE**

The recommendations to improve the S-CAPE are organized as follows: changes that should be made to the S-CAPE itself, the information students receive from the assessment and how, after improvements are made, that the S-CAPE could be used to control placement. The reliability of the S-CAPE needs to be increased. This can be done by “eliminating items with low item discrimination values and increasing the number of items that a student sees. Furthermore, changing the CAPE algorithm so that items are not chosen solely on item difficulty, but rather require students to answer all of the item types (grammar, vocabulary, and reading)” (Cox, 2014, personal communication). In other words, first some of the items need to be eliminated or updated based on their ability to differentiate between the students’ levels. The S-CAPE is unidimensional, meaning that it puts grammar, vocabulary, and reading comprehension knowledge all on the same axis. However, students will not have equal knowledge in each area. For example, some students may have differential vocabulary
development. To eliminate this problem, the test questions could be grouped in testlets. Instead of moving up or down levels based on the result of one question, the test could present all three test types at the same level.

As related to the author, many professors and instructors of other universities acknowledge BYU for creating a placement test that is widely used, but upon further questioning they acknowledge that its questions are out of date. For example, Figure 5.4 is one of the questions taken from the S-CAPE.

Figure 5.4. Example S-CAPE Question That Needs Updating.

The author, after working in the public school system, feels that this question does not reflect the current culture of religious neutrality in education. Although BYU is a private, religious institution, the S-CAPE is being used by many other institutions, which further supports why the S-CAPE needs to be updated. Additionally, the S-CAPE has fewer items at higher levels of difficulty. If students take the S-CAPE just once, the flagging system of the S-CAPE (see page 18) will prevent them from seeing an item more than once. However, if students take the S-
CAPE more than once, they may see the same questions, and that, in turn, could affect the outcome of the test. Therefore, items that are more difficult need to be included.

Also, the current components of grammar, vocabulary, and reading comprehension only represent a part of proficiency. When the S-CAPE was first created, its original author recommended that a listening component be added. Now, 28 years later, the researcher of this study agrees. An oral component should also be added, such as Elicited Oral Response tests. These recommendations are based on the belief that the Department of Spanish and Portuguese at BYU would like to align its courses with an overarching goal of helping students develop communicative skills. By adding them, the construct validity of the S-CAPE for the BYU context would be strengthened.

Students need to know what their S-CAPE results mean. If the S-CAPE is to stay as it is, student need to be advised in the actual S-CAPE instructions to take the exam twice. Currently, students only receive this recommendation from instructors or department secretaries who are familiar with the test. One student stated in their survey comments, “I took it last year, 2012, to help determine whether or not I should be in 105 or 101/102. I found it very helpful once I took it multiple times. I averaged my score out of the 3 times I took it. If I had only taken it once, I do not think it would have been very accurate.” One weakness of this recommendation is that it is not back by statistics. In addition, after the exam, instead of receiving just one class recommendation, students could be shown how they might perform in two different courses. This is to help them understand their results better and to better meet their needs.

Also, to better meet students’ needs, the author recommends the updated cutoff scores that appear in Table 5. These recommended cutoff scores were based on contrasting group statistics. The cutoff scores between classes were computed by starting with the means of the test
scores of two sequential classes. The difference between the two means was then divided by two. Lastly, that number was added to the lower one of the cutoff scores. For example, the cutoff scores for SPAN 105 (mean 422) and SPAN 106 (mean 432) was 427, and it was obtained through this formula: (432-422)/2 = 5. 422 + 5= 427. Table 5 provides the current placement recommendation that BYU uses for students’ S-CAPE results compared to the recommended placement scores based on this research study. They are needed because, as was stated previously, the current cutoff scores were originally based on students’ tests scores and then were modified when four Spanish classes (SPAN 101, 102, 201, 202) were turned into six Spanish classes (SPAN 101, 102, 105, 106, 205 and 206).

Table 5

<table>
<thead>
<tr>
<th>Current Placement Scores</th>
<th>Recommended Placement Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPAN 101</td>
<td>Below 276</td>
</tr>
<tr>
<td>SPAN 102</td>
<td>276-315</td>
</tr>
<tr>
<td>SPAN 105</td>
<td>316-370</td>
</tr>
<tr>
<td>SPAN 106</td>
<td>(No specific recommendation for SPAN 106)</td>
</tr>
<tr>
<td>SPAN 205</td>
<td>371-420</td>
</tr>
<tr>
<td>SPAN 206</td>
<td>421-520</td>
</tr>
<tr>
<td>SPAN 321</td>
<td>Above 520</td>
</tr>
</tbody>
</table>

Notice that the difficulty continuum of the recommended placement scores is higher. (Since, this study did not include SPAN 101 and 102, no cutoff scores could be recommended for those
classes.) The author views the recommended cutoff scores as an improvement because they are based on students’ test results instead of just expert opinion, and therefore, they are more descriptive of students enrolled in the SPAN 105–321 courses. They are also noticeably higher than the previous recommendations, which will encourage students to be more prepared for each class. Nevertheless, there is some degree of similarity in the ranges between the current and the recommended scores.

There are three weaknesses in these recommendations. First of all, participants took the S-CAPE at the beginning of the semester instead of before the semester began, which may have affected the results because students may have learned some new material at that point. Also, it is hard to say if the students that took the S-CAPE were truly representative of their levels due to the fact that they placed themselves at the beginning of the semester. Lastly, the instructors who reported on how well the S-CAPE placed their students submitted their feedback based on the current placement chart used by the Department of Spanish and Portuguese, which means that they may have responded differently if the recommended cutoff scores had been used.

After these improvements are implemented, requiring students to take the S-CAPE in order to register for class could better control the placement process. The benefit of implementing such a requirement is eliminating the students’ assumptions of what classes they need to take, especially for returned missionaries because, even though SPAN 321 is known as the class that they should take, it may not be meeting the students’ needs. Another benefit is eliminating the beginning of semester changes where students jump from one class to another. The author believes that many students are lost in this process of placement. Once students see that the current class they are taking is not right for them, they often feel that they cannot change to another class because other classes are already full at the beginning of the semester. Classes
end up not being full as students drop out or change classes during the first week. Hence, an S-CAPE controlled placement system would eliminate some of the extra movement at the beginning of the semester.

These recommendations for the improvement of the S-CAPE show us that this test is part of a larger placement system. The purpose of the following suggestions is to improve that system.

**Pedagogical Implications**

In analyzing the placement system of the BYU Department of Spanish and Portuguese, the author recommends the following: updating course outcomes and analyzing and explicitly stating placement expectations.

Course outcomes for SPAN 101–321 (including conversation courses) are vague in some cases and lacking articulation between courses. The author believes that in order to better meet students’ needs and place students, that what students do in each course needs to be better defined. For example, the following are the course outcomes for SPAN 206.

- “Consolidate the student’s command of the basic structures of the Spanish language.
- “Develop literacy and oral language skills at the intermediate-advanced range.
- “Obtain an adequate knowledge of the Spanish language and cultures through the study of selected perceptions and manifestations.
- “Understanding the natural speech of native speakers from any Spanish-speaking country and interpret the general idea in their proper cultural background” (Undergraduate Catalogue, 2014).

Notice how it would be difficult for students to understand what they would be learning in this class. Figure 5.5 shows the analysis of these course outcomes (Undergraduate Catalogue, 2014)
using an online resource entitled Wordle. This web resource controls font size by how frequently a word appears in a given source: the more frequently a word appears the larger it will be.

Figure 5.5. Analysis of Learning Outcomes for SPAN 101–206 Using Wordle.net.

From this visual, we can see that the main focus of the courses is “Spanish”, which is where it should be. Another word that stands out because of its size is “understand” and “understanding”. Aside from the fact that these words represent one of the lowest levels of Bloom’s Taxonomy, they do not stress the importance of communication. The Spanish courses seem to struggle between the focus of grammar (as connected to the word “understand”) and oral skills. On the main webpage for the Spanish department, it is stated, “Based on the liberal arts tradition, the Spanish programs at BYU are designed not only to improve students’ communication skills but to immerse them in a study of Hispanic linguistics, literature and culture” (Department of Spanish and Portuguese, 2014). Additionally, the Program Learning Outcomes stated that students will be able to, “Develop the four major language skills in Spanish (speaking, reading,
writing and listening comprehension) in a variety of formal and informal situations, narrating and describing in all time frames and successfully handling linguistically unfamiliar situations” (Department of Spanish and Portuguese, 2014). There is a lack of articulation between the course outcomes, the overall program outcomes and what is taught in the classroom. The importance of communication is stated, but not always practiced. The placement process will be better if what is expected in each course and how courses will build on each is more clearly articulated. This needs to be done before the S-CAPE can be updated, to ensure, as was emphasized earlier, the importance having a placement exam that matches the placement context in which it is used.

After course outcomes have been clearly defined, they can act as a table of test specifications for the updated S-CAPE. These specifications can make oral skills part of what is being tested, instead of the placement of students solely on grammar, vocabulary, and reading comprehension knowledge. This could possibly be based on OPI ratings. ACTFL, in conjunction with ACE (American Council on Teaching), made recommendations of college course equivalents for official OPI ratings (American Council on Teaching, 2014). This is more effective than using years studied as an equivalent to a college course due to the fact that the number of years a student studies does not guarantee a certain level of proficiency. This is the exact opposite of BYU’s placement method, but is more aligned with BYU’s overall goal in Spanish proficiency development.

In addition to an increased focus on communication, course outcomes could be based on the Can-Do Statements recently created by NCSSFL-ACTFL (see Appendix E). This would also allow students to place themselves based on the information provided in the course outcomes. NCSSFL-ACTFL claims, “The more learners are engaged in their own learning process, the more intrinsically motivated they become. Research shows that the ability of language learners to
set goals is linked to increased student motivation, language achievement, and growth in proficiency” (NCSSFL-ACTFL, 2014, p. 1). The author recommends BYU create similar statements for its students, or adopt those created by NCSSFL-ACTFL. This kind of placement has shown to be beneficial because if students are involved in selecting their own class, there is a greater chance that they will stay with that class (LeBlanc and Painchaud, 1985). This is in contrast to the many changes students currently make between Spanish classes at the beginning of each school semesters as they try to figure out which class they should be in. Some students already see the need for this resource. In the survey given, one student [referring to the S-CAPE] said, “It has some gaps in my opinion. I think better descriptions of classes, for example, writing samples of the students each class is intended for, this would allow students to self-place based on their fellow students would be extremely helpful.” This form of educated self-assessment would enable BYU’s Spanish courses to be proficiency based, and also establish the importance of proficiency from introductory courses through the course sequence. Students would then be able to track their progress throughout their BYU experience. It would also enable them to continue studying on their own since they will be better educated about their own language abilities.

Finally, the BYU Spanish and Portuguese Department needs to track these students’ progress to see if they are meeting course and program outcomes. A continual survey to assess the type of background students taking basic Spanish classes have also needs to be implemented. This will help the BYU Department of Spanish and Portuguese monitor the changes in demographic in their student population.
Suggestions for Future Research

Future research needs to look more at how intermediate and advanced language students could be placed. Eggington and Cox (2014) suggested that Elicited Oral Response (EOR) could be used for language placement. However, future research needs to be done to actually turn EOR into a placement test. In addition to oral testing, writing skills and their assessment also needs to be made more feasible.

Conclusions

There seems to be greater face validity to the S-CAPE among instructors and students than the researcher anticipated, even though the reliability of the S-CAPE is not high enough to give users great confidence in the instrument (.64). However, the S-CAPE seems to be a good resource to be used in conjunction with other means of placement.

Students tend to rely on the limited information presented in the course catalog listed when they register for a class. The author recommends that the BYU Department of Spanish and Portuguese begin by revising the information in the course catalog to give students a better overview of the course. This could be an extension of what is available on the BYU Learning Outcomes website. That website could include a link to the Spanish courses on the registration page so that students, along with their S-CAPE score, can decide which course best meet their needs.

In addition to the limited information available on the BYU Learning Outcomes website, survey results show that students get a lot of their information used in making registration decisions from talking to family or friends. Therefore, in addition to updating the Learning Outcomes, the author recommends that a brief PowerPoint of the results of this study be presented on the first day of class with the Learning Outcomes so that students not only know
themselves, but also begin to educate each other regarding the classes for which they should register. Giving students the tools to place themselves more effectively would be one strategy that students could use to bridge the gap.
Appendix A – S-CAPE Testing Information Included in Course Syllabuses

“BYU is conducting a research study. The purpose of this study is to better place students studying Spanish into an appropriate language class. As part of this study, you will be taking the S-CAPE (Spanish Computerized Adaptive Placement Exam) during the weeks of September 16th–27th. Since it is part of a research study, this test must be taken in the HTRSC Testing Lab (B151 JFSB), and the rules and schedules of the testing lab apply. Unlike other tests, you will not be able to sign up for a time to take this test, which means you need to go early. You will receive credit for taking this test if completed twice (no partial credit will be given). The test needs to be taken twice for reliability purposes, which means that you can take it twice in a row or on two different occasions. Your participation is voluntary, meaning that even though you are required to take the test for credit, you will be given the option, before the test begins, of not having your results included in the research study. All test results will be kept confidential. Please see your instructor for further information.”
Appendix B – Survey Questions Included with the S-CAPE Exam

1. What is your student ID number?

2. Please select your age

3. Please select your gender

4. Please select the current Spanish class you are taking

5. Select how many years you have been studying Spanish

   3 months, 6 months, 1 year, 2 years, 3 years, 4 years, 5 years, 6 years, 7 years

6. Select how you learned Spanish (select all that apply and select for how long you were involved):

   a) elementary school
   b) heritage speaker (Spanish was spoken in your home)
   c) high school
   d) immersion experience
   e) middle school/junior high
   f) missionary service
   g) native speaker (Fluent in Spanish)
   h) self-instructed
   i) study abroad
   j) university classes

[The following data was lost due to a computer error, except for #10.]

7. If you served a mission, was it Spanish-speaking? [yes/no] Where? ________ How long?

   3 months, 6 months, 9 months, 12 months, 18 months, 24 months

8. Select all that apply:

   a) I took this placement test before registering for classes this semester
   b) I registered for the class recommended by the placement test
   c) I have taken the placement test in previous semesters

9. On what did you base your decision to take this course (select all that apply):
a) Course description  

b) Faculty recommendation  

c) Former Spanish teacher  

d) Information from BYU Spanish and Portuguese website  

10. Will you allow your testing results to be included in the study of this test?
Appendix C – Information Included in Emails to Instructors

We would like to ask how well you felt the S-CAPE placed your students. Your feedback will help us to better update the S-CAPE exam. However, you are not required to participate.

On the following scale, please rate how many students this test correctly placed:

10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Do you have any other feedback about the S-CAPE you would like to include?

Current placement recommendations based on S-CAPE scores, which ranges are listed in the left-hand column:

| below 276 | Spanish 101 |
| 276–315 | SPAN 101/102 accelerated |
| 316–370 | SPAN 101/102 accelerated or SPAN 105 |
| 371–420 | SPAN 105/106 accelerated or SPAN 205 |
| 421–520 | SPAN 206 |
| above 520 | SPAN 321 |
Appendix D – Information Included in Emails to Students

Students who received an email had taken a Spanish course between SPAN105–321 during Fall 2013.

Dear Spanish Student,

At the beginning of this semester, many of you took a placement exam called the S-CAPE for course credit. We have a few follow up questions we would ask you to answer. Even if you did not take the S-CAPE, you are welcome to participate. These questions will help us to know how to better place students in the future. It will take you about 3–4 minutes to answer. Thank you for your participation.

Survey Questions:

1. Please select the Spanish class you took Fall 2013
   a. SPAN 105
   b. SPAN 106
   c. SPAN 205
   d. SPAN 206
   e. SPAN 321

2. How many semesters have you been studying at a university level?
   a. 1 semester
   b. 2 semesters
   c. 3 semesters
   d. 4 semesters
   e. 5 semesters
   f. 6 semesters
   g. Other: ____________

3. Did you serve a mission for The Church of Jesus Christ of Latter-day Saints?
   a. Yes
   b. No

4. If you selected “yes”, where did you serve? (Please list the name of the country)
   a. _____________________
5. How long were you on your mission? (Please list in months. Example: 6 months)
   a. ____________________

6. What percentage of your mission did you speak Spanish?
   a. 100%  
   b. 90%  
   c. 80%  
   d. 70%  
   e. 60%  
   f. 50%  
   g. 40%  
   h. 30%  
   i. 20%  
   j. 10%

7. Did you take the S-CAPE at the beginning of Fall Semester 2013? (Not as part of a class assignment, but to help you determine what class you should take)
   a. Yes  
   b. No

8. Did you follow its recommendations?
   a. Yes  
   b. No

9. Did you feel like the S-CAPE correctly placed you?
   a. Yes  
   b. No

10. How did you choose your Spanish class?
    a. By its course description
    b. It was the next course in the sequence
    c. Teacher/professor recommendation
    d. Other: ____________________

11. How do you feel about the S-CAPE as a placement tool?
    ____________________
Appendix E – Can-Do Statements

“For Learners: How can you use the Can-Do Statements? Use the checklist to record what you think you can do. You may realize that your progress may not be the same for each mode: Interpersonal, Interpretive, or Presentational. This is to be completely expected. For example, you may progress more quickly in Interpretive Reading than in Interpersonal Communication. You will begin to determine your progress on the proficiency ladder by assessing each mode separately.

“For Learning Facilitators: How can you use the Can-Do Statements? Use the Global Can-Do Benchmarks and main indicators to set long-term learning goals. Ask yourself what you expect your learners to be able to do with language after one semester, after one year, or after several years, and re-evaluate your goals when informed by assessment. Choose more specific can-do statements or customize new ones to establish learning targets for thematic units and lessons. These targets can help drive your instruction to be more performance-oriented and provide more opportunities for your learners to produce language” (NCSSFL-ACTFL, 2014, p. 1-2).
<table>
<thead>
<tr>
<th><strong>Advanced Mid</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interpersonal Communication</strong></td>
</tr>
<tr>
<td>I can express myself fully not only on familiar topics but also on some concrete social, academic, and professional topics. I can talk in detail and in an organized way about events and experiences in various time frames. I can confidently handle routine situations with an unexpected complication. I can share my point of view in discussions on some complex issues.</td>
</tr>
<tr>
<td><strong>Presentational Speaking</strong></td>
</tr>
<tr>
<td>I can deliver well-organized presentations on concrete social, academic, and professional topics. I can present detailed information about events and experiences in various time frames.</td>
</tr>
<tr>
<td><strong>Presentational Writing</strong></td>
</tr>
<tr>
<td>I can write on a wide variety of general interest, professional, and academic topics. I can write well-organized, detailed paragraphs in various time frames.</td>
</tr>
<tr>
<td><strong>Interpretive Listening</strong></td>
</tr>
<tr>
<td>I can understand the main idea and most supporting details on a variety of topics of personal and general interest, as well as some topics of professional interest. I can follow stories and descriptions of some length and in various time frames. I can understand information presented in most genres, even when not familiar with the topic.</td>
</tr>
<tr>
<td><strong>Interpretive Reading</strong></td>
</tr>
<tr>
<td>I can understand the main idea and most supporting details in texts on a variety of topics of personal and general interest, as well as some professional topics. I can follow stories and descriptions of considerable length and in various time frames. I can understand texts written in a variety of genres, even when I am unfamiliar with the topic.</td>
</tr>
</tbody>
</table>
Interpersonal Communication

ADVANCED MID
I can express myself fully not only on familiar topics but also on some concrete social, academic, and professional topics. I can talk in detail and in an organized way about events and experiences in various time frames. I can confidently handle routine situations with an unexpected complication. I can share my point of view in discussions on some complex issues.

- I can communicate effectively on a wide variety of present, past, and future events.
- I can give a clear and detailed story about childhood memories, such as what happened during vacations or memorable events and answer questions about my story.
- I can give detailed descriptions about cultural events and respond to questions about them.
- I can talk about present challenges in my school or work life, such as paying for classes or dealing with difficult colleagues.
- I can discuss future plans, such as where I want to live and what I will be doing in the next few years.

Presentational Speaking

ADVANCED MID
I can deliver well-organized presentations on concrete social, academic, and professional topics. I can present detailed information about events and experiences in various time frames.

- I can present information about events of public or personal interest.
- I can recount the details of a historical event.
- I can present in detail the plot, setting, characters, etc. of a film or book.
- I can describe in detail a social event or a local celebration.
- I can present a full account of the social and cultural activities from a recent trip or excursion.
- I can tell a story to a particular audience for dramatic effect.
- I can incorporate simple analogies into presentations.
- I can give an accurate description of something I participated in or witnessed.

- I can give presentations with ease and detail on a wide variety of topics related to professional interests.
- I can give a presentation about my studies, work, or organization to an outside audience.
- I can advocate for new ideas or innovative approaches related to school, work, or training.
- I can present detailed information to clients, customers, or others.

- I can convey my ideas and elaborate on a variety of academic topics.
- I can make presentations on a variety of subjects I have researched.
- I can teach a lesson intended for a particular audience.
- I can give detailed presentations on the process and the outcome of an experiment, research study, etc.
- I can make presentations to advocate for educational opportunities such as membership in a club, honor society, or study-abroad.
- I can give a presentation on a capstone or similar summative project such as a thesis or seminar.
Presentational Writing

ADVANCED MID
I can write on a wide variety of general interest, professional, and academic topics. I can write well-organized, detailed paragraphs in various time frames.

I can write well organized texts for a variety of academic purposes.
- I can write a research paper on a topic related to my studies or area of specialization.
- I can write a proposal for a project or a research study.
- I can write a newspaper and/or magazine article about an event, project, or research initiative.
- I can write content for instructional resources.
- I can ________________

I can write well organized texts for a variety of professional purposes.
- I can write letters of recommendation.
- I can write a project proposal or a report.
- I can write about the results of a survey and the recommendations that might follow.
- I can write a statement for a job application.
- I can write performance reviews or project evaluations.
- I can write brochures or other resources for clients or customers.
- I can ____________________

Interpretive Reading

ADVANCED MID
I can understand the main idea and most supporting details in texts on a variety of topics of personal and general interest, as well as some professional topics. I can follow stories and descriptions of considerable length and in various time frames. I can understand texts written in a variety of genres, even when I am unfamiliar with the topic.

I can follow the general idea and some details of what is written in a variety of stories and autobiographical accounts.
- I can follow the chronological plot in a simple short story.
- I can follow news articles reporting on community events.
- I can follow a short online autobiography.
- I can follow a cover letter and a resume.
- I can ____________________

I can understand general information on topics outside my field of interest.
- I can understand the details about a police report on a recent crime.
- I can understand the details of a job or performance evaluation.
- I can understand the details of an article about a sporting or cultural event.
- I can ____________________
REFERENCES


Eggington, W. & Cox, T. (2013). Using elicited oral response testing to measure speaking and
listening proficiency in order to determine the need for an interpreter. *Harvard Latino Law Review* 16, 127-146.


